Spatial Statistics 14 (2015) 303–320

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Signal-Controlled Least Square Collocation: A new quality in geostatistical estimation and simulation?

Joachim Menz [a,1], Jörg Benndorf [b,*], Tom Wambeke [b,c]

[a] *TU Bergakademie Freiberg, Institut für Markscheidewesen und Geodäsie, Reiche Zeche, Fuchsmühlenweg 9, D 09599 Freiberg, Germany*

[b] *Delft University of Technology, Department of Geosciences and Engineering, Stevinweg 1, 2628 CN Delft, P.O. Box 5048, 2600 GA Delft, The Netherlands*

[c] *IHC Merwede, Smitweg 6, 2961 AW, P.O. Box 8, 2960AA Kinderdijk, The Netherlands*

A R T I C L E   I N F O

Article history:
Received 8 November 2014
Accepted 23 June 2015
Available online 30 June 2015

Keywords:
Cross-validation
Fitting model parameters
Collocation
Dual Kriging

A B S T R A C T

Geostatistical estimation and simulation algorithms are designed to provide the most likely forecast as well as information about the accuracy of the prediction. The representativeness of these measures strongly depends on the quality of inferred model parameters, which are mainly defined by the parameters of the variogram or the covariance function. Available data may be scarce, trend affected and of different data type making the inference of representative geostatistical model parameters difficult.

This contribution introduces a new method for best fitting of geostatistical model parameters in the presence of a trend, which utilizes empirical and theoretical differences between Universal Kriging and trend-predictions. The method is called "Signal-Controlled Collocation" and extends well known approaches of cross validation in two aspects. Firstly, the model evaluation is not only limited to sample data locations but is performed on any prediction location of the attribute in the domain. Secondly, it extends the measure used in cross validation, based on a single point replacement by using error curves. These allow the definition of rings of influence representing errors resulting from separate variogram lags.

* Corresponding author. Tel.: +31 15278 6024.
E-mail addresses: menz@mabb.tu-freiberg.de (J. Menz), j.benndorf@tudelft.nl (J. Benndorf), t.wambeke@tudelft.nl (T. Wambeke).

[1] Tel.: +49 03731 392606.

http://dx.doi.org/10.1016/j.spasta.2015.06.008

The formal description of the method is provided followed by two case studies related to the prediction of mining-induced ground movements and resource modelling of heavy mineral deposits. Recommendations for future research are given.

1. Introduction

Geostatistical estimation and simulation algorithms are designed to provide the most likely forecast and information about the prediction error. This is a main advantage over other spatial interpolation techniques and makes them highly applicable for assessing spatial uncertainty and risk. The representativeness of geostatistical accuracy measures strongly depends on the quality of the inferred model parameters. These are mainly defined by the parameters of the variogram or the covariance function. With the aim of an objective derivation of optimal model parameters this contribution proposes an extension to the traditional cross-validation (Stone, 1974; Dubrule, 1983), the *Signal-Controlled Collocation* (Menz, 2012a,b). The complete theoretical background of the method was first presented in Benndorf and Menz (2014) and extends the approaches of cross-validation in two aspects. Firstly, the model evaluation is not only limited to sample data locations. It can be performed on any prediction location of the attribute in the domain. Secondly, it extends the measure used in cross validation, based on single point replacements, by using error curves. These relate the prediction error to different lag distances and allow to distinguish the influence of several model parameters on the model performance. A tailored back-fitting procedure permits the derivation of model parameters that fit best the data available under the assumption of a given functional trend model.

Signal-Controlled Collocation utilizes mean errors of empirical and theoretical differences between Universal Kriging (UK) (Matheron, 1969) and a trend-prediction by general least squares GLS. Errors are calculated for a set of m grid nodes as a function of the distance between the data and grid nodes to estimate. The grid contains in total M grid nodes and covers the complete domain to estimate or simulate. Dual Kriging (Matheron, 1981) is used for predicting the attribute under consideration $Z(\boldsymbol{x})$, which is composed of a signal component $S(\boldsymbol{x})$, a trend $m(\boldsymbol{x})$ and a noise component $R(\boldsymbol{x})$

$$Z(\boldsymbol{x}) = S(\boldsymbol{x}) + m(\boldsymbol{x}) + R(\boldsymbol{x}). \tag{1}$$

Note that in geodesy this approach is well known as Least-Square Collocation (Krarup, 1969; Moritz, 1970, 1973, 1989), which is the simultaneous performance of adjustment, prediction and filtering. It is worth mentioning that both approaches, Dual Kriging and Least-Square Collocation, are equivalent. The derivation of the formalism in geodetic adjustment theory and geostatistics differs substantially.

The mean error of the empirical and theoretical differences, the signal $\hat{S}(\boldsymbol{x})$

$$\hat{S}(\boldsymbol{x}) = \hat{Z}(\boldsymbol{x})_{UK} - \hat{m}(\boldsymbol{x})_{GLS}, \tag{2}$$

is depicted in error-curves as a function of the distance between data and location to estimate. These curves provide the means to adjust the covariance model parameters by back-fitting. This can be achieved by an incremental variation of the model parameters until the theoretical error curve matches the empirical one (Fig. 1). Vertical deviations between the two error curves are corrected by varying the variances σ_S^2 and σ_R^2 of the model components signal and noise. Horizontal deviations are resolved by adjusting the range a, which is the range of spatial dependence of the stationary signal. In this way the influence of individual model parameters on the prediction error can be distinguished.

The basis of the proposed method is formed by empirical and theoretical mean errors of the signal components, which are calculated by Least-Square Collocation. This is reflected in the name of the method, *Signal-Controlled Collocation*.